

## The Theta model

Konstantinos Galanos

The theta model involves the decomposition of a time series into a set of newly calculated series that contain less information.

The coefficient  $\theta$  is applied directly to second differences of the time series:

$$X''_{new}(\theta) = \theta \cdot X''_{data} \quad (1)$$

where

$$X''_{data} = X_t - 2X_{t-1} + X_{t-2} \text{ at time } t$$

If:

- $\theta = 0$             the time series is transformed into a linear regression line
- $0 < \theta < 1$         the time series is deflated
- $\theta > 1$              the time series is dilated

The general formulation of the Theta method is:

1. Initial time series is decomposed into two or more Theta lines.
2. Each of the Theta lines is extrapolated separately and the forecasts are simply combined with equal weights.

Different combinations of Theta lines can be employed for each forecasting horizon.

### **Example**

In the special case of a two theta lines decomposition ( $\theta = 0, \theta = 2$ ) the following formula stands:

$$X_t = \frac{1}{2}[L_t(\theta = 0) + L_t(\theta = 2)], \forall t = 1, \dots, n \quad (1)$$

The theta line for  $\theta = 0$  is a LRL (Linear Regression Line). Therefore, in order to easily calculate the  $\theta = 2$  theta line as follows:

$$\begin{aligned} X_t &= \frac{1}{2}[LRL_t + L_t(\theta = 2)] \\ \Rightarrow L_t(\theta = 2) &= 2X_t - LRL_t \end{aligned} \quad (2)$$

One may take the following steps to execute a forecast using this specific theta method:

- 1) Apply Linear Regression to data and prepare LRL and forecasts
- 2) Prepare values for  $L_t(\theta = 2)$  using formula (2)
- 3) Extrapolate  $L(\theta = 2)$  with either SES (Simple Exponential Smoothing) or other smoothing method
- 4) Combine with equal weights the forecasts from SES and LR (Linear Regression)

### **Formula of the Theta Point Forecast**

$$F_{t+1} = \alpha Y_t + (1 - \alpha)F_t + \frac{\hat{b}}{2} \quad (3)$$

where

$\alpha$	level smoothing parameter
$F_{t+1}$	point forecast
$\hat{b}$	linear regression slope
$Y_t$	data point

The co-originators of the Theta model (K. Nikolopoulos and V. Assimakopoulos) are arguing that the SES+d method proposed by Hyndman as a simple version of the former model is actually what it is stated to be. They state that it is based on a particular model determined using the methodology of the Theta model, in essence replicating a single condition.

They also state that another difference between the two models is the smoothing parameter  $\alpha$ . In the Theta model this is calculated via a Mean Square Error (MSE) minimization procedure based on the Theta line  $L_t(\theta=2)$  time series. In the case of the SES+d model,  $\alpha$  is minimized with regard to the original data time series.

### **REFERENCES**

1. **“The theta model: a decomposition approach to forecasting”**, V. Assimakopoulos, K. Nikolopoulos, Department of Electrical and Computer Engineering, Forecasting Systems Unit, National Technical University of Athens, 15773 Zografou, Athens, Greece